

# Some remarks on the staircasing phenomenon in total-variation based image denoising

Khalid Jalalzai\*

April 10, 2012

## Abstract

This paper deals with the so-called staircasing phenomenon, which frequently arises in total variation based denoising models in image analysis. We prove in particular that staircasing always occurs at global extrema of the datum and at all extrema of the minimizer. It is also shown that, for radial images, the staircasing always appears at the extrema and at the boundary of the image. We also prove the equivalence between the denoising model and the total variation flow, in the radial case, thus extending a previous result in dimension one. This equivalence cannot hold in the nonradial case, as it is shown with a counterexample. This connection allows us to understand how the staircase zones and the discontinuities of the denoising problem evolve with the regularization parameter.

**Keywords:** total variation, image denoising, staircasing.

**AMS:** 35J70, 65J20, 35K65, 68U10.

**Acknowledgment:** I warmly thank Antonin Chambolle who suggested me this fruitful research project.

## 1 Introduction

Functions of bounded variation equipped with the total variation semi-norm were introduced for image reconstruction in 1992. Since then, they have had many successful applications for inverse problems in imaging. Indeed, the penalization of the total variation has the ability to smooth out the image by creating large regular zones and to keep the edges of the most important objects in the image. In this paper we aim to study the latter property in the continuous setting.

We assume that a corrupted image  $g : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$  went through a degradation

$$g = g_0 + n$$

where  $g_0$  is the original clean image,  $n$  is a Gaussian white noise of standard deviation  $\sigma$ . Rudin, Osher and Fatemi (ROF) proposed in [29] to minimize the total variation

$$u \mapsto TV(u) = \int_{\Omega} |Du|$$

---

\*CMAP, CNRS, Ecole Polytechnique, 91128 Palaiseau Cedex, France  
(khalid.jalalzai@polytechnique.edu).

amongst functions of bounded variation under the constraint  $\|u - g\|_2^2 \leq \sigma^2 |\Omega|^2$  to solve the inverse problem and thus get a restored image  $u$ . It was proven in [17] that one can solve in an equivalent way the unconstrained problem

$$\min_{u \in BV(\Omega)} \lambda \int_{\Omega} |Du| + \frac{1}{2} \|u - g\|_2^2$$

for an adequate Lagrange multiplier  $\lambda$ . In the literature the minimization of ROF's energy is referred to as the *denoising problem*.

Another possibility is to consider the *total variation flow* for restoring  $g$ . A denoised image is given by  $u(t)$  that solves

$$\begin{cases} -\partial_t u(t) \in \partial TV(u(t)) \text{ a.e. } t \in [0, T], \\ u(0) = g. \end{cases}$$

which has a unique solution according to [5]. As we shall see it is not true in general that these approaches coincide.

It has been long observed that using the total variation has the advantage of recovering the discontinuities quite well. In [12], the authors carried out a study of the behavior of the minimizer of the denoising problem at these discontinuities. Their results are extended in [23, 24] where it is proven that the discontinuities of the minimizer of the anisotropic total variation are contained in those of the datum  $g$ . Moreover, it is established that one can observe new discontinuities in the weighted case if the weight is merely Lipschitz continuous.

In the present paper, we shall focus on another very important property of the total variation: it smoothes the highly oscillating regions by creating large constant zones which is known in the literature as the *staircasing effect*. This phenomenon is sometimes not desirable in imaging applications since it yields blocky and non-natural structures. It was already studied in [28] for the one-dimensional case. Actually, the author proves that whenever the data  $g \notin BV(a, b)$ , the minimizer  $u_\lambda$  vanishes almost everywhere. In [27], Nikolova proves that the staircasing effect is related to the non-differentiability of the total variation term. More precisely, large homogeneous zones are recovered from noisy data and remain unchanged for small perturbations. In other words, the creation of such zones is quite probable. On the contrary, absence of staircasing with the differentiable approximation  $TV_\varepsilon = \sqrt{|Du|^2 + \varepsilon}$  is also proven. The latter approaches were carried out only for finite dimensional approximations of the total variation. In [26], Louchet and Moisan proposed an alternative to the minimization of the total variation by considering the TV-LSE filter (see [25] for another alternative). In [26], the authors proved in the discrete setting that TV-LSE avoids the staircasing effect (in the sense that a region made of 2 pixels or more where the restored image is constant almost never occurs). As far as we know, there is no result on the subject in the higher dimensional and continuous setting for the classical total variation functional. We shall show that staircasing always occurs (even without addition of noise) at global extrema of the datum and at all extrema of the minimizer.

In the last section, we investigate further these qualitative properties. An interesting question is to understand how the staircase zones and the discontinuities evolve with the regularization parameter  $\lambda$ . The idea is to use the results that are already established for the flow. Unfortunately, in higher dimension the connection between the flow and ROF's energy fails (we give a counterexample). However, we are going to prove that this connection actually holds for radial functions. This way, one can prove that the discontinuities form a decreasing sequence, whereas the staircase zones increase with the regularization parameter (which is not true in general).

Let us remark that all the results are established in dimension  $N \geq 2$  since the situation is quite well understood in the one-dimensional case and was widely studied in the literature (see

the recent paper [7] for instance). Indeed, in dimension one, ROF's denoising problem reads as follows

$$\min_{u \in BV(\mathbb{R})} \int \lambda |u'(x)| + \frac{1}{2}(u - g)^2(x) dx \quad (1.1)$$

for  $g \in L^2(\mathbb{R})$  and some positive real  $\lambda$ . Let us denote  $u_\lambda$  the minimizer of this problem.

Writing down the Euler-Lagrange (see equation 2.1) one immediately sees that either  $u_\lambda$  is constant or  $z_\lambda = \text{sgn}(u'_\lambda)$  and as a consequence  $u_\lambda = g$ . This is an almost explicit formulation of the solution that tells us that

- the discontinuities of  $u_\lambda$  are contained in those of  $g$ ,
- flat zones are created at maxima and minima of  $g$ .

This can be seen in the following simulation:

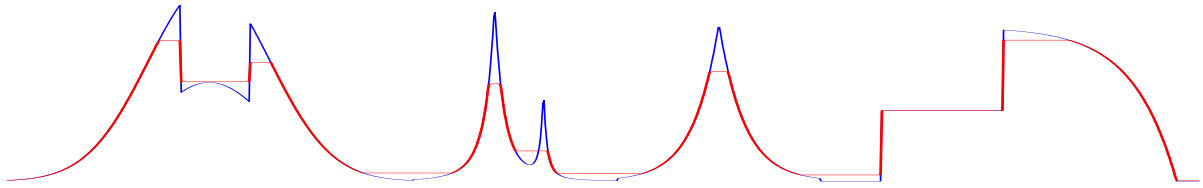


Figure 1: Minimizer  $u_\lambda$  (in red) of a 1D data  $g$  (in blue).

The other important result in dimension one is the link with the flow of the total variation. It is known that problem (1.1) with  $\lambda = t$  is minimized by  $u(t)$ , the unique solution of the flow. In the recent article [10], the authors used this relation to prove that for a signal that went through an addition of noise (that is the trajectory of a Wiener process strictly speaking), staircasing occurs almost everywhere. This observation seems to be more general as can be seen in the following test:

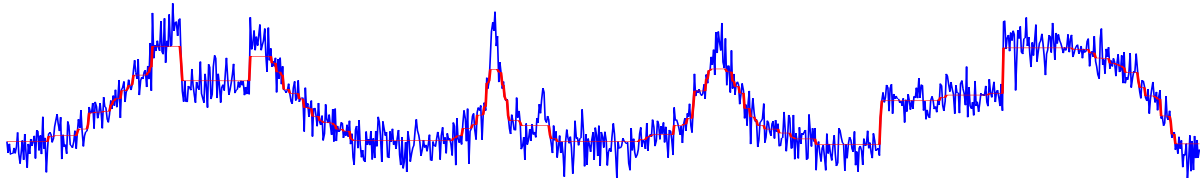


Figure 2: Minimizer  $u_\lambda$  (in red) of a noisy data  $g$  (in blue) is constant almost everywhere.

## 2 Mathematical preliminary

Henceforth  $\Omega$  will denote an open subset of  $\mathbb{R}^N$  with Lipschitz continuous boundary. The material of this section can be found in the classical textbooks [4, 21, 31] but also in the recent survey [16].

### 2.1 Functions of bounded variation

Let us start with following fundamental definition:

**Definition 2.1.** A function  $u \in L^1(\Omega)$  is of bounded variation in  $\Omega$  (denoted  $u \in BV(\Omega)$ ) if its distributional derivative  $Du$  is a vector-valued Radon measure that has finite total variation i.e.  $|Du|(\Omega) < \infty$ . By the Riesz representation theorem, this is equivalent to say that

$$|Du|(\Omega) = \sup \left\{ \int_{\Omega} u \operatorname{div} \varphi \mid \varphi \in C_c^\infty(\Omega, \mathbb{R}^N), \forall x \in \Omega \ |\varphi(x)| \leq 1 \right\} < \infty.$$

In the sequel, the quantity  $|Du|(\Omega)$  also denoted  $\int_{\Omega} |Du|$  or simply  $TV(u)$  will be called the total variation of  $u$ . It is readily checked that  $\|\cdot\|_1 + TV$  defines a norm on  $BV(\Omega)$  that makes it a Banach space.

A first result that is a straightforward consequence of the dual definition we just gave is a key step to apply the direct method in the calculus of variations:

**Proposition 2.2** (Sequential lower semicontinuity). Let  $(u_n)_{n \in \mathbb{N}}$  be any sequence in  $BV(\Omega)$  such that  $u_n \rightarrow u$  in  $L^1(\Omega)$  then

$$\int_{\Omega} |Du| \leq \liminf_{n \rightarrow \infty} \int_{\Omega} |Du_n|.$$

One also has

**Proposition 2.3** (Approximation by smooth functions). If  $u \in BV(\Omega)$  then there exists a sequence  $(u_n)_{n \in \mathbb{N}}$  of functions in  $C^\infty(\Omega)$  such that

$$\begin{aligned} u_n &\rightarrow u \text{ in } L^1(\Omega) \\ \int_{\Omega} |\nabla u_n| &\rightarrow \int_{\Omega} |Du|. \end{aligned}$$

For the direct method to apply it is usually of importance to have a compactness result:

**Theorem 2.4** (Rellich's compactness in  $BV$ ). Given a bounded  $\Omega \subset \mathbb{R}^N$  with Lipschitz boundary and any sequence  $(u_n)_{n \in \mathbb{N}}$  such that  $\left( \|u_n\|_{L^1(\Omega)} + \int_{\Omega} |Du_n| \right)$  is bounded, there exists a subsequence  $(u_{n(k)})_{k \in \mathbb{N}}$  that converges in  $L^1$  to some  $u \in BV(\Omega)$  as  $k \rightarrow \infty$ .

**Definition 2.5.** Let  $E \subset \mathbb{R}^N$  be a Borelian set. It is called a set of finite perimeter or also Cacciopoli set if  $u = \chi_E$  is a function of bounded variation. We will call perimeter of  $E$  in  $\Omega$ , and denote  $P(E, \Omega)$  or simply  $P(E)$  (if  $\Omega = \mathbb{R}^N$ ), its total variation.

The following key result provides a connection between the total variation of a function and the perimeter of its level sets.

**Theorem 2.6** (Coarea formula). If  $u \in BV(\Omega)$ , the set  $E_t = \{u > t\}$  has finite perimeter for a.e.  $t \in \mathbb{R}$  and

$$|Du|(B) = \int_{-\infty}^{\infty} |D\chi_{\{u>t\}}|(B) dt$$

for any Borel set  $B \subset \Omega$ .

Finally let us recall the following

**Theorem 2.7** (Sobolev inequalities). *Let  $u \in BV(\Omega)$  and let us denote  $\langle u \rangle = \frac{1}{|\Omega|} \int_{\Omega} u$ . For a bounded Lipschitz domain  $\Omega$ , the following Poincaré inequality holds*

$$\|u - \langle u \rangle\|_{L^{\frac{N}{N-1}}(\Omega)} \leq C(N, \Omega) \int_{\Omega} |Du|.$$

*If  $\Omega = \mathbb{R}^N$ , one has the Sobolev inequality*

$$\|u\|_{L^{\frac{N}{N-1}}(\mathbb{R}^N)} \leq C(N) \int_{\Omega} |Du|.$$

*In particular, if  $u = \chi_E$ , one gets the isoperimetric inequality*

$$|E|^{\frac{N-1}{N}} \leq C(N) P(E, \Omega).$$

## 2.2 BV functions in image processing

The classical model of a functional where total variation plays a key role is the so-called Rudin-Osher-Fatemi energy:

$$\mathcal{E}_{\lambda}(u) = \lambda \int_{\Omega} |Du| + \frac{1}{2} \|u - g\|_2^2 \quad (\text{ROF})$$

In the sequel we shall be interested in minimizing this energy in  $BV(\Omega)$  for some positive real  $\lambda$ . By Proposition 2.2, there is a minimizer in  $BV(\Omega)$ , denoted  $u_{\lambda}$  in the sequel (uniqueness follows from the strict convexity of the energy).

The energy  $\mathcal{E}_{\lambda}$  is not smooth though convex so it is still possible to get the Euler-Lagrange equation as follows:

**Proposition 2.8.** *Function  $u_{\lambda}$  minimizes  $\mathcal{E}_{\lambda}$  in  $BV(\Omega)$  if and only if there exists  $z_{\lambda} \in L^{\infty}(\Omega, \mathbb{R}^N)$  such that*

$$\begin{cases} -\lambda \operatorname{div} z_{\lambda} + u_{\lambda} = g & \text{in } \Omega, \\ |z_{\lambda}| \leq 1 & \text{in } \Omega, \\ z_{\lambda} \cdot Du_{\lambda} = |Du_{\lambda}|, & \\ z_{\lambda} \cdot \nu = 0 & \text{on } \partial\Omega, \end{cases} \quad (2.1)$$

*with  $\nu$  denoting the inner normal to  $\Omega$ .*

In the sequel, we shall only consider Neumann boundary conditions but we could of course take into account Dirichlet or more complicated conditions.

**Remark 2.9.**  $z_{\lambda} \cdot Du_{\lambda}$  is the pairing of a bounded function with a bounded measure and should be understood in the sense of Anzelotti [6].

By the coarea formula, the superlevel sets  $\{u_{\lambda} > t\}$  are sets of finite perimeter for almost every  $t$  that satisfy the following minimal surface problem:

**Theorem 2.10.** *Let  $u_{\lambda}$  be the minimizer of (ROF). Then for any  $t \in \mathbb{R}$ ,  $\{u_{\lambda} > t\}$  (resp.  $\{u_{\lambda} \geq t\}$ ) is the minimal (resp. maximal) solution of the minimal surface problem*

$$\min_E \lambda P(E, \Omega) + \int_E (t - g(x)) dx \quad (2.2)$$

*over all sets of finite perimeter in  $\Omega$ . Moreover  $\{u_{\lambda} > t\}$  being defined up to negligible sets, there exists an open representative.*

### 3 Staircasing for the denoising problem

As recalled in our introduction, it has been long observed that the minimizer of this problem has unnatural homogeneous regions referred to as the *staircase* regions. The problem of proving the existence of these constant zones has been tackled in the discrete setting in [27] but almost nothing is known in the continuous setting. We will prove in this section that the staircasing phenomenon is unavoidable in the continuous setting in dimension  $N \geq 2$  even though there is no addition of noise.

Staircasing through the level lines:

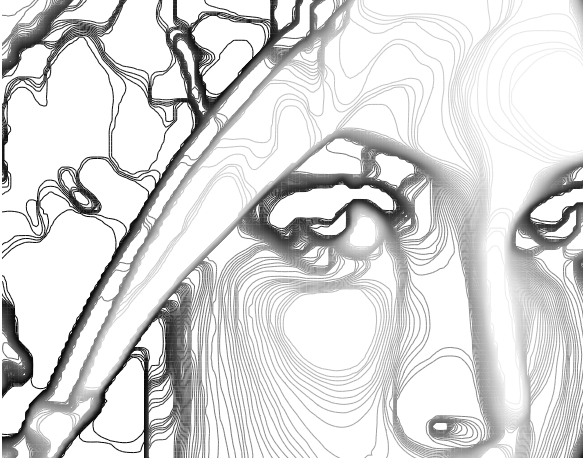


Figure 3: Level lines for the TV-minimizer of the *Lena* image

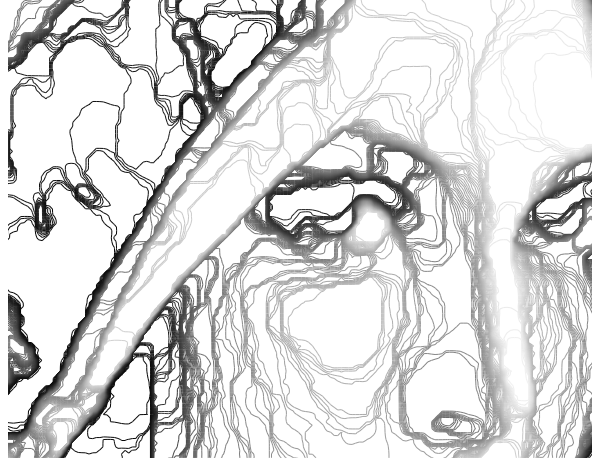


Figure 4: Level lines for the TV-minimizer of a noisy image

The previous images show that the level lines of the minimizers miss large regions that are therefore constant. The idea of looking at the level sets gives a good intuition but it is also a key point in our analysis to establish results from a theoretical point of view.

#### 3.1 Staircasing at extrema

Let us start by stating one of the main results of this section:

**Theorem 3.1.** *Let  $g \in L^2(\mathbb{R}^N)$  bounded from above. Then the associated minimizer  $u_\lambda$  of (ROF) ( $\lambda > 0$ ) is also bounded from above, attains its maximum and one has*

$$|\{u_\lambda = \max u_\lambda\}| > 0.$$

*In particular  $Du_\lambda = 0$  in  $\{u_\lambda = \max u_\lambda\}$ . A similar result holds for the minimum when  $g$  is bounded from below.*

*Proof.* We denote  $m_g := \sup g$ . Proving that there is a staircase amounts to show that the superlevel sets vanish at some point. Let us therefore consider the superlevel  $E_t^\lambda := \{u_\lambda > t\}$  for some real  $t$ . By Theorem 2.10, we know that  $E_t^\lambda$  minimizes the following problem

$$\min_E \lambda P(E) + \int_E (t - g(x)) dx.$$

In particular,

$$\lambda P(E_t^\lambda) + \int_{E_t^\lambda} (t - g(x)) dx \leq 0. \quad (3.1)$$

By the isoperimetric inequality (see Theorem 2.7) and equation (3.1), we get

$$|E_t^\lambda|^{\frac{N-1}{N}} \leq P(E_t^\lambda) \leq \frac{1}{\lambda} \int_{E_t^\lambda} (g(x) - t) dx \leq |E_t^\lambda| \frac{m_g - t}{\lambda}.$$

This implies immediately that  $|E_t^\lambda| = 0$  for any  $t \geq m_g$ . Actually, by a thresholding argument, one can show that  $E_t^\lambda = \emptyset$ , but this additional information will not be needed here.

If for some  $t \in \mathbb{R}$ ,  $|E_t^\lambda| \neq 0$  then

$$|E_t^\lambda| \geq \left( \frac{\lambda}{m_g - t} \right)^N. \quad (3.2)$$

Let us assume that we have  $|E_t^\lambda| \neq 0$  for any  $t_0 < t < m_g$ . Then letting  $t \rightarrow m_g$  in (3.2) contradicts  $|E_t^\lambda| < +\infty$ . So, if we set  $m_u := \sup u_\lambda$ , we therefore have  $m_u = \sup\{t, |E_t^\lambda| \neq 0\} < m_g$ .

Now, we would like to prove that  $|\{u_\lambda = m_u\}| \neq 0$ . This is indeed true since

$$\{u_\lambda = m_u\} = \bigcap_{n \in \mathbb{N}} E_{m_u - \frac{1}{n}}^\lambda$$

and by (3.2) it follows

$$|\{u_\lambda = m_u\}| \geq \lim_{n \rightarrow \infty} |E_{m_u - \frac{1}{n}}^\lambda| \geq \left( \frac{\lambda}{m_g - m_u} \right)^N > 0.$$

By the coarea formula, we then simply have

$$\begin{aligned} \int_{\{u_\lambda = m_u\}} |Du_\lambda| &= \int_{-\infty}^{+\infty} P(E_t^\lambda \cap \{u_\lambda = m_u\}) dt \\ &= \int_{m_u}^{m_g} P(E_t^\lambda) dt \\ &= 0. \end{aligned}$$

This implies that  $Du_\lambda = 0$  on the staircase set  $\{u_\lambda = m_u\}$  that has positive Lebesgue measure.  $\square$

**Remark 3.2.**(i) In case  $g \in L^\infty(\mathbb{R}^N)$  is not assumed to be constant, we actually proved that  $u$  is also bounded and that we have

$$\inf_{\mathbb{R}^N} g < \min_{\mathbb{R}^N} u_\lambda \leq \max_{\mathbb{R}^N} u_\lambda < \sup_{\mathbb{R}^N} g.$$

Moreover, inequality (3.2) gives a lower bound for the staircasing effect namely

$$\begin{aligned} |\{Du_\lambda = 0\}| &\geq |\{u_\lambda = \min_{\mathbb{R}^N} u_\lambda\} \cup \{u_\lambda = \max_{\mathbb{R}^N} u_\lambda\}| \\ &\geq 2 \left( \frac{\lambda}{\sup_{\mathbb{R}^N} g - \inf_{\mathbb{R}^N} g} \right)^N. \end{aligned} \quad (3.3)$$



(ii) Our proof is rather simple but as was stated in Theorem 2.10, it is possible to prove that  $\{u_\lambda = \max_\Omega u_\lambda\}$  has an open representative. However, the proof is not direct and relies on the well-known density estimate for quasi-minimizers of the perimeter (see Lemma 3.4 and the proof of Theorem 3.3).

As we just saw, Theorem 3.1 furnishes a way to quantify the staircase effect through the inequality (3.3). Nonetheless, this bound is not sharp as can be seen for  $g$  the characteristic of a convex set (see below). The reason is that we do not take into account creation of flat zones occurring near local extrema. This is the object of the following theorem

**Theorem 3.3.** *Let  $g \in L^p(\Omega)$  with  $p \in (N, +\infty]$  and  $u_\lambda$  with  $\lambda > 0$  the corresponding minimizer of (ROF). If  $x_0$  is a local maximum of  $u_\lambda$  then there exists  $C > 0$  depending only on  $N$  such that*

$$\liminf_{r \rightarrow 0} \frac{|\{u_\lambda = u_\lambda^+(x_0)\} \cap B(x_0, r)|}{|B(x_0, r)|} > C.$$

In particular, for any radius  $\rho > 0$ ,

$$|\{u_\lambda = u_\lambda^+(x_0)\} \cap B(x_0, \rho)| > 0.$$

Moreover, if  $x_0 \notin J_{u_\lambda}$  there exists  $\mathcal{N}(x_0)$  a neighborhood of  $x_0$  such that

$$Du_\lambda = 0 \text{ on } \mathcal{N}(x_0).$$

A similiar result holds for a local minimum.

To prove the theorem we shall need the following result:

**Lemma 3.4** (Density estimate). *Consider  $E$  a set of finite perimeter that is a minimizer of the perimeter problem (2.2). Then, denoting  $\omega_N$  the volume of the unit ball in  $\mathbb{R}^N$ , there exists a radius  $r_0 > 0$  such that for any point  $x \in \Omega$ ,*

- if  $\forall r > 0, |E \cap B(x, r)| > 0$  then  $\forall r < r_0, |E \cap B(x, r)| \geq \frac{\omega_N r^N}{2^N}$ ,
- if  $\forall r > 0, |B(x, r) \setminus E| > 0$  then  $\forall r < r_0, |B(x, r) \setminus E| \geq \frac{\omega_N r^N}{2^N}$ .

In particular, if  $x \in \partial^* E$ ,

$$\forall r < r_0, \min(|E \cap B(x, r)|, |B(x, r) \setminus E|) \geq \frac{\omega_N r^N}{2^N}.$$

One may find a proof of this classical result in [24, Lemma 5.1] and also in [11, 13, 19]. With the previous lemma in hands we can now turn to the proof of Theorem 3.3:

*Proof.* Without loss of generality, we can consider a local maximum point  $x_0$  of  $u_\lambda$  that is to say there is a radius  $\rho > 0$  such that

$$u_\lambda \leq u_\lambda^+(x_0) \text{ on } B(x_0, \rho).$$



Considering that  $x_0 \in E_t^\lambda = \{u_\lambda > t\}$  for some  $t < u_\lambda^+(x_0)$ , one has for any  $r > 0$ ,

$$|E_t^\lambda \cap B(x_0, r)| > 0,$$

and by the Density Lemma 3.4

$$\liminf_{r \rightarrow 0} \frac{|E_t^\lambda \cap B(x_0, r)|}{\omega_N r^N} \geq \frac{1}{2^N}.$$

Now given that for any  $r > 0$ ,

$$|\{u_\lambda = u_\lambda^+(x_0)\} \cap B(x_0, r)| = \inf_{\varepsilon > 0} |E_{u_\lambda^+(x_0) - \varepsilon}^\lambda \cap B(x_0, r)|$$

then by the Lebesgue's density theorem,

$$\liminf_{r \rightarrow 0} \frac{|\{u_\lambda = u_\lambda^+(x_0)\} \cap B(x_0, r)|}{\omega_N r^N} = \inf_{\varepsilon > 0} \liminf_{r \rightarrow 0} \frac{|E_{u_\lambda^+(x_0) - \varepsilon}^\lambda \cap B(x_0, r)|}{\omega_N r^N} \geq \frac{1}{2^N}.$$

So there is some small  $r < \rho$  such that

$$|\{u_\lambda = u_\lambda^+(x_0)\} \cap B(x_0, \rho)| \geq \frac{\omega_N r^N}{2^N}.$$

Moreover, if we identify  $\{u_\lambda = u_\lambda^+(x_0)\}$  with the set of points where it has density one, then it is an open set.  $\square$

Putting together these two theorems we get

**Corollary 3.5.** *Let  $g \in L^2(\mathbb{R}^N) \cap L^\infty(\mathbb{R}^N)$  and  $u_\lambda$ ,  $\lambda > 0$  be the minimizer of (ROF) associated to  $g$ . If  $u_\lambda$  is constant on at most two disjoint sets then it has no local extrema other than its global maximum and minimum.*

### 3.2 Dependency of the staircasing on $\lambda$

In the previous section, we proved that for fixed  $\lambda$  staircasing always occurs and can be quantified by (3.1). This bound suggests that the Lebesgue measure of the created flat zones is non-decreasing with respect to  $\lambda$ . This was already observed for the total variation flow in  $\mathbb{R}^N$ . In [5, Chapter 4] the authors even prove that the solution  $u(t)$  of the total variation flow in  $\mathbb{R}^N$  decreases in time, for some norm, with a finite extinction time. It is possible to get a similar result for the minimizer  $u_\lambda$ :

**Proposition 3.6.** *Let  $\Omega$  be a connected bounded Lipschitz continuous open set in  $\mathbb{R}^N$ ,  $g \in L^N(\Omega)$  and  $u_\lambda$  the minimizer of (ROF). Then, there exists  $\lambda^* = C_\Omega \|g\|_N \geq 0$  with  $C_\Omega$  that only depends on  $\Omega$ , such that for any  $\lambda \geq \lambda^*$ ,*

$$u_\lambda = \frac{1}{|\Omega|} \int_\Omega g.$$

*Proof.* The conclusion follows readily if one can find some  $z_\lambda$  that satisfies the system

$$\begin{cases} -\lambda \operatorname{div} z_\lambda + \frac{1}{|\Omega|} \int_\Omega g = g & \text{in } \Omega, \\ \|z_\lambda\|_\infty \leq 1 & \text{in } \Omega, \\ z_\lambda \cdot \nu_\Omega = 0 & \text{on } \partial\Omega. \end{cases}$$

If  $p \geq N$ , the function  $g - \frac{1}{|\Omega|} \int_{\Omega} g$  is in  $L^N(\Omega)$  and of mean zero. Therefore, the result of Bourgain and Brezis [8] (see also [20]) asserts that there exists a  $z \in L^\infty(\Omega, \mathbb{R}^N) \cap W^{1,N}(\Omega, \mathbb{R}^N)$  that solves

$$\begin{cases} -\operatorname{div} z = g - \frac{1}{|\Omega|} \int_{\Omega} g & \text{in } \Omega, \\ z \cdot \nu_{\Omega} = 0 & \text{on } \partial\Omega. \end{cases}$$

Thus a nice candidate is  $z_{\lambda} = \frac{z}{\lambda}$ , for  $\lambda$  large enough. Actually,  $\lambda^* = \|z\|_{\infty}$  that is controlled by  $\|g\|_N$ .  $\square$

If the domain is not bounded, the result of Bourgain and Brezis does not apply. However, one has a similar result for a data  $g$  that lies in the so-called *Schwartz class*  $\mathcal{S}$  that contains those functions whose derivatives are decreasing faster than any polynomial (see [22]):

**Proposition 3.7.** *If  $g$  is an element of  $\mathcal{S}(\mathbb{R}^N)$  and  $u_{\lambda}$  denotes the minimizer of (ROF) corresponding to  $g$  then there exists  $\lambda^* \geq 0$  such that for  $\lambda \geq \lambda^*$*

$$u_{\lambda} = 0 \text{ in } \mathbb{R}^N.$$

*Proof.* As in the previous proof, the assertion follows if one can find some  $z_{\lambda}$  that solves the following system

$$\begin{cases} -\lambda \operatorname{div} z_{\lambda} = g, \\ \|z_{\lambda}\|_{\infty} \leq 1. \end{cases}$$

Let us look for a  $z_{\lambda}$  that is of the form  $z_{\lambda} = \frac{\nabla v}{\lambda}$  with  $v$  that satisfies in  $\mathbb{R}^N$

$$-\Delta v = g. \tag{3.4}$$

Taking the Fourier transform on both sides, one gets

$$-4\pi^2 |\xi|^2 \hat{v}(\xi) = \hat{g}(\xi), \quad \forall \xi \in \mathbb{R}^N$$

hence the following estimate

$$\|\nabla v\|_{\infty} \leq \|\widehat{\nabla v}\|_1 = \frac{1}{4\pi^2} \| |\xi|^{-1} \hat{g}(\xi) \|_1 < +\infty$$

where in the last inequality we used the well-known fact that  $\hat{g} \in \mathcal{S} \subset L^p(\mathbb{R}^N)$  for any  $p \in [1, \infty]$  (see [22] for instance for further details). Therefore  $z_{\lambda} = \frac{\nabla v}{\lambda}$  satisfies the system above as soon as  $\lambda \geq \lambda^* := \|\nabla v\|_{\infty}$ .  $\square$

For a general unbounded subdomain of  $\mathbb{R}^N$ , the previous proof cannot be adapted since it is well known that for a bounded  $g$ , equation (3.4) does not necessarily admit a solution  $v$  in  $W^{2,\infty}$  (see [8]).

In case  $N = 2$ , it is possible to prove Proposition 3.6 without having to use the difficult result of Bourgain and Brezis:

**Proposition 3.8.** *Let  $\Omega \subset \mathbb{R}^2$  be a connected open set that is bounded with a Lipschitz continuous boundary,  $g \in L^2(\Omega)$  and  $u_\lambda$  the minimizer of (ROF) associated to  $g$ . Then there exists  $\lambda^* = C_\Omega \|g\|_2$ , with  $C_\Omega$  that only depends on  $\Omega$ , such that for  $\lambda \geq \lambda^*$*

$$u_\lambda = \frac{1}{|\Omega|} \int_\Omega g.$$

*Proof.* Given  $u \in BV(\Omega) \cap L^2(\Omega)$ , in [5, Lemma 2.4], one can find the following characterization of  $p = -\operatorname{div}(z) \in \partial TV(u)$ :

$$\int_\Omega |Du| \leq \int_\Omega (u - \varphi)p + \int_\Omega z \cdot \nabla \varphi, \quad \forall \varphi \in W^{1,1}(\Omega) \cap L^2(\Omega).$$

Therefore, denoting  $\langle u_\lambda \rangle$  the average of  $u_\lambda$  and setting  $u = u_\lambda$ ,  $\varphi = \langle u_\lambda \rangle$  and  $p = \frac{1}{\lambda}(g - u_\lambda)$  one has

$$\int_\Omega |Du_\lambda| \leq \frac{1}{\lambda} \int_\Omega (g - u_\lambda)(u_\lambda - \langle u_\lambda \rangle).$$

Then, applying the Poincaré inequality (see Theorem 2.7) and Cauchy-Schwarz and using  $\langle u_\lambda \rangle = \langle g \rangle$ , the average of  $g$ , one obtains the estimate

$$C \|u_\lambda - \langle g \rangle\|_2 \leq \frac{1}{\lambda} \|u_\lambda - \langle g \rangle\|_2 \|u_\lambda - g\|_2,$$

where  $C$  is the constant that appears in the Poincaré inequality. Now remarking that, by minimality of  $u_\lambda$ ,

$$\|u_\lambda - g\|_2 \leq \|g\|_2$$

concludes the proof. Moreover, we get  $\lambda^* = \frac{\|g\|_2}{C}$ . □

By the Sobolev inequality (see Theorem 2.7), for which the optimal constant is known, it is readily checked that we also gave an alternative proof for Proposition 3.7 in case  $N = 2$ :

**Proposition 3.9.** *Let  $g \in L^2(\mathbb{R}^2)$  and  $u_\lambda$  be the minimizer of (ROF) associated to  $g$  then there exists  $\lambda^* = \left(2\pi^{\frac{1}{2}}\right)^{-1} \|g\|_2$  such that for  $\lambda \geq \lambda^*$*

$$u_\lambda = 0 \text{ in } \mathbb{R}^N.$$

**Remark 3.10.** *Reasoning as we did we get an explicit  $\lambda^*$  that is optimal as can be seen for  $g = \chi_{\mathbb{D}}$  the characteristic of the unit disc since we know in this case that*

$$u_\lambda = (1 - 2\lambda)^+ \chi_{\mathbb{D}}.$$

Now that we got rid of the case when  $\lambda$  is large, let us see through some examples how the staircase regions behave for reasonable values of the regularization parameter. In the rest of this section, we assume that  $g$  is the characteristic function of a set  $C$  which means we are now interested in the minimizers of

$$\mathcal{E}_\lambda(u) = \lambda \int_\Omega |Du| + \frac{1}{2} \|u - \chi_C\|_2^2.$$

We are going to distinguish two different cases for  $C$ :

### 3.2.1 The characteristic of a bounded convex set in $\mathbb{R}^2$

Then it is well known (see [1,3]) that for  $\lambda \leq \lambda^*$  and for any  $t \geq 0$  the superlevel  $E_t^\lambda = \{u_\lambda > t\}$  of the solution  $u_\lambda$  is given by

$$E_t^\lambda = \begin{cases} C_{\lambda/(1-t)} & \text{if } t \leq 1 - \lambda/R^*, \\ \emptyset & \text{otherwise.} \end{cases}$$

where for any  $R > 0$ ,  $C_R$  is the opening of  $C$  defined by

$$C_R = \bigcup_{B(x,R) \subset C} B(x,R),$$

and  $R^*$  is the inverse of the so-called *Cheeger constant* defined by the value of  $R$  that solves

$$\frac{P(C_R)}{|C_R|} = \frac{1}{R}.$$

Therefore, the staircase set

$$\{u_\lambda = \max u_\lambda\} = C_{R^*}$$

is the so-called *Cheeger set* and is independent of  $\lambda \leq \lambda^*$ . We would get similar results if  $C$  were the union of spaced convex sets (see [5] for the expression of  $u_\lambda$  in this case).

### 3.2.2 The characteristic of two touching squares in $\mathbb{R}^2$

Here  $C = [0,1] \times [0,-1] \cup [-1,0] \times [0,1]$  is the union of two unit squares that only touch on a vertex. In [2], Allard gives a full description of the superlevels of the solution  $u_\lambda$ , hence  $u_\lambda$  itself. The level sets  $E_t^\lambda = \{u_\lambda > t\}$  are of five kinds namely

$$E_t^\lambda \in \left\{ \emptyset, \mathbb{R}^2, F_{r,s}, G_{r,s}, H_r \mid \lambda = \frac{1}{r} + \frac{1}{s}, r, s \in \mathbb{R}^+ \right\}$$

where the last three sets are formally defined in [2]. They are depicted in the following figures as the interior of the domain bounded by the red curve:

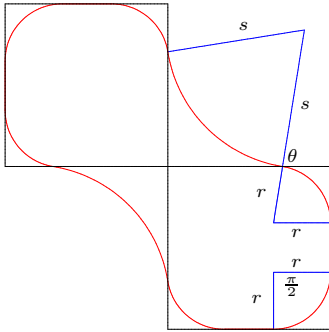


Figure 5:  $F_{r,s}$ .

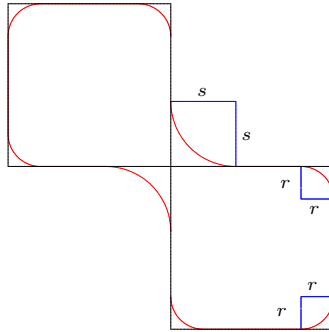


Figure 6:  $G_{r,s}$ .

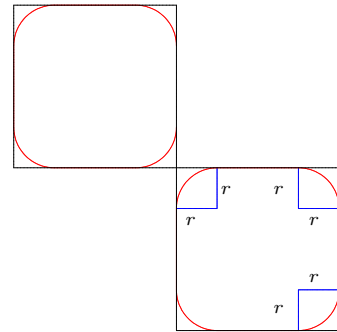


Figure 7:  $H_r$ .

In the following figure Allard summed up the different possibilities:

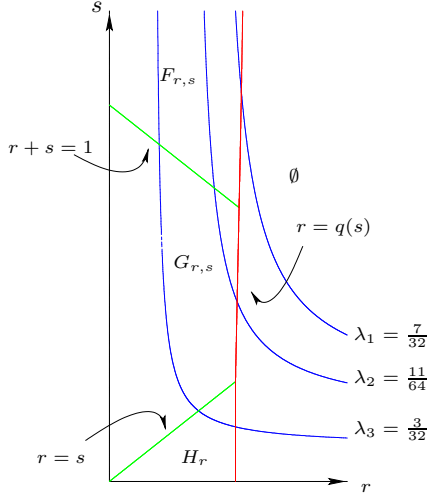


Figure 8: Level sets  $E_t^\lambda$ .

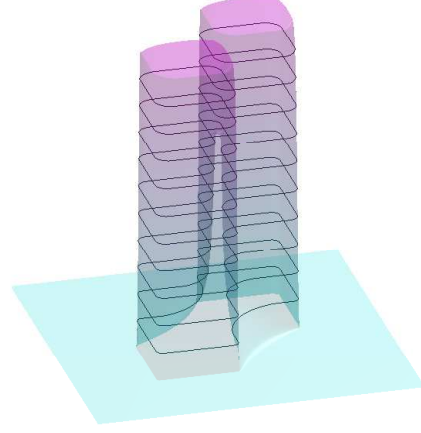


Figure 9:  $u_\lambda$  and its level lines.

Now consider that  $\lambda \in (0, \lambda^*)$ . As seen previously, the staircase set  $S_\lambda := \{u_\lambda = \max u_\lambda\}$  is exactly the smallest superlevel set that is not empty this is to say

$$S_\lambda \in \{F_{q(s),s}, G_{q(s),s}, H_{q(s)}\}$$

where the function  $q(s)$  is non-decreasing in  $s$ . Let us focus on two values of the regularization parameter:  $\lambda_1 = \frac{7}{32}$  and  $\lambda_2 = \frac{11}{64}$  and let  $s_i$ ,  $i \in \{1, 2\}$  be the unique value such that  $\lambda_i = \frac{1}{s_i} + \frac{1}{q(s_i)}$ . Then it is readily seen that

$$S_{\lambda_2} = G_{q(s_2),s_2} \not\subset F_{q(s_1),s_1} = S_{\lambda_1}$$

even though  $\lambda_2 < \lambda_1$ .

This example shows that in general the staircase zones  $(S_\lambda)_{\lambda \geq 0}$  do not form a monotone sequence.

Such a phenomenon cannot occur given a radial function  $g$  and this is what we are going to prove in the following section. In short, this is due to the fact that the solutions  $(u_\lambda)_\lambda$  of the radial problem form a semi-group. This was already established in the one-dimensional setting [10] and in case  $g = \chi_C$  the characteristic function of a convex set  $C$  [3, 5].

## 4 Denoising problem for radial data

Unless otherwise stated, in this section  $\Omega$  will denote the ball  $B(0, R) \subset \mathbb{R}^N$  with  $N \geq 2$  and we consider a radial  $g \in L^2(\Omega)$ . It is easily seen that the minimizer  $u_\lambda$  of (ROF) is itself radial. Indeed one could argue that for any rotation  $\mathcal{R}$ ,  $u_\lambda(\mathcal{R}x)$  is also a minimizer. Since it is unique  $u_\lambda(\mathcal{R}x) = u_\lambda(x)$  for any  $x \in \mathbb{R}^N$ . We denote  $\tilde{v}$  the function defined by  $\tilde{v}(|x|) = v(x)$  for any  $x \in \mathbb{R}^N$ . Then  $u_\lambda$  minimizes

$$\min_{u \in BV(\Omega)} \int_0^R \left( \lambda |\tilde{u}'(r)| + \frac{1}{2} (\tilde{u}(r) - \tilde{g}(r))^2 \right) r^{N-1} dr. \quad (4.1)$$

Thus, proceeding as in the one-dimensional case, either  $u_\lambda$  is constant or

$$u_\lambda(x) = g(x) + \operatorname{sgn}(\tilde{u}'_\lambda(|x|)) \frac{(N-1)\lambda}{|x|}$$

for any  $x \in \mathbb{R}^N \setminus \{0\}$ . Now, introducing the dual variable  $\tilde{z}$  and reasoning as in [15], we can derive a dual formulation for the minimization problem (4.1), that is

$$\inf_{\substack{\tilde{z} \in W^{1,2}(0,R), \\ \tilde{z}(R)=0, |\tilde{z}| \leq \lambda}} \int_0^R \left( \tilde{z}'(r) + \frac{N-1}{r} \tilde{z}(r) + g(r) \right)^2 r^{N-1} dr.$$

Then, if we set  $z(x) = \tilde{z}(x) \frac{x}{|x|}$  for any  $x \in \mathbb{R}^N \setminus \{0\}$ ,

$$(\operatorname{div}(z))(x) = \nabla(\tilde{z}(|x|)) \cdot \frac{x}{|x|} + \tilde{z}(|x|) \operatorname{div} \left( \frac{x}{|x|} \right) = \tilde{z}'(|x|) + \tilde{z}(|x|) \frac{N-1}{|x|}$$

which gives the dual formulation

$$\inf_{z \in A_\lambda} \mathcal{E}(z) := \int_{B(0,R)} (\operatorname{div}(z) + g)^2, \quad (4.2)$$

where

$$A_\lambda = \left\{ z(x) = \tilde{z}(x) \frac{x}{|x|} / \forall x \in \mathbb{R}^N \setminus \{0\}, \tilde{z} \in W^{1,2}(0,R), \tilde{z}(R) = 0, |\tilde{z}| \leq \lambda \right\}$$

is the set of admissible  $z$ . Henceforth, we denote  $z_\lambda$  a minimizer of problem (4.2). Since it is radial, we will sometimes write  $z_\lambda$  instead of  $\tilde{z}_\lambda$ . Note also that  $z_\lambda$  is actually continuous by the Sobolev embedding theorem.

**Remark 4.1.** *We decided to proceed this way to justify rigorously that one can pick a radial vectorfield  $z_\lambda$  in the Euler-Lagrange equation (2.1).*

#### 4.1 The solutions form a semi-group

Let us start with the following lemma:

**Lemma 4.2.** *Let  $g \in L^p(\Omega)$  with  $p \in [N, +\infty]$  then for any  $\lambda > 0$ ,*

$$|z_\lambda(x)| \leq C|x|^{1-\frac{N}{p}} \|g - u_\lambda\|_{L^p(B(0,|x|))}$$

for some positive  $C$  hence in particular  $z_\lambda(0) = 0$ .

*Proof.* For a.e.  $r = |x|$  with  $x \in \mathbb{R}^N$

$$\operatorname{div}(z_\lambda)(x) = \tilde{z}'_\lambda(r) + \frac{N-1}{r} \tilde{z}_\lambda(r) = \frac{1}{r^{N-1}} (\tilde{z}_\lambda(r) r^{N-1})'.$$

Hence integrating with respect to  $r$  and using  $\operatorname{div}(z_\lambda) \leq |g - u_\lambda|$ , it follows

$$\tilde{z}_\lambda(r) r^{N-1} \leq \int_{B(0,r)} |g - u_\lambda| \leq C \|g - u_\lambda\|_{L^p(B(0,r))} r^{N(1-\frac{1}{p})}$$

for some positive real  $C$ . □

The following proposition is the key result in the study of the radial problem:

**Proposition 4.3** (Comparison result). *Let  $g \in L^2(\Omega)$ ,  $\mu > \lambda \geq 0$  and consider respectively  $z_\lambda$  and  $z_\mu$  the corresponding minimizers of  $\mathcal{E}$  then one has*

$$z_\lambda - \lambda \geq z_\mu - \mu.$$

*Proof.* On the one hand,

$$\begin{aligned} \mathcal{E} \left( \left( z_\mu + (\lambda - \mu) \frac{x}{|x|} \right) \vee z_\lambda \right) &= \int_{\{z_\lambda > z_\mu + \lambda - \mu\}} (\operatorname{div}(z_\lambda) + g)^2 dx \\ &\quad + \int_{\{z_\lambda < z_\mu + \lambda - \mu\}} \left( \operatorname{div} \left( z_\mu + (\lambda - \mu) \frac{x}{|x|} \right) + g \right)^2 dx, \\ \mathcal{E} \left( \left( z_\lambda + (\mu - \lambda) \frac{x}{|x|} \right) \wedge z_\mu \right) &= \int_{\{z_\lambda > z_\mu + \lambda - \mu\}} (\operatorname{div}(z_\mu) + g)^2 dx \\ &\quad + \int_{\{z_\lambda < z_\mu + \lambda - \mu\}} \left( \operatorname{div} \left( z_\lambda + (\mu - \lambda) \frac{x}{|x|} \right) + g \right)^2 dx, \end{aligned}$$

thus by minimality of  $z_\mu$  and  $z_\lambda$

$$\begin{aligned} &\int_{\{z_\lambda < z_\mu + \lambda - \mu\}} \left( (\operatorname{div}(z_\lambda) + g)^2 + (\operatorname{div}(z_\mu) + g)^2 \right) dx \\ &\leq \int_{\{z_\lambda < z_\mu + \lambda - \mu\}} \left( \operatorname{div} \left( z_\mu + (\lambda - \mu) \frac{x}{|x|} \right) + g \right)^2 dx \\ &\quad + \int_{\{z_\lambda < z_\mu + \lambda - \mu\}} \left( \operatorname{div} \left( z_\lambda + (\mu - \lambda) \frac{x}{|x|} \right) + g \right)^2 dx, \end{aligned}$$

and then it follows that

$$\int_{\{z_\lambda - z_\mu + \mu - \lambda < 0\}} (\mu - \lambda) \frac{N-1}{|x|} \left( \operatorname{div} \left( z_\lambda - z_\mu + (\mu - \lambda) \frac{x}{|x|} \right) \right) dx \geq 0. \quad (4.3)$$

But on the other hand, given that  $z_\lambda - z_\mu + \mu - \lambda > 0$  on  $\partial\Omega$  and  $z_\lambda - z_\mu + \mu - \lambda = 0$  on  $\partial\{z_\lambda - z_\mu + \mu - \lambda < 0\}$ , one gets by integration for  $\varepsilon > 0$ ,

$$\begin{aligned} &\int_{[\Omega \setminus B(0, \varepsilon)] \cap \{z_\lambda - z_\mu + \mu - \lambda < 0\}} \frac{N-1}{|x|} \left( \operatorname{div} \left( z_\lambda - z_\mu + (\mu - \lambda) \frac{x}{|x|} \right) \right) dx \\ &= -\frac{N-1}{\varepsilon} \int_{\partial B(0, \varepsilon) \cap \{z_\lambda - z_\mu + \mu - \lambda < 0\}} (z_\lambda - z_\mu + \mu - \lambda)(x) dx \\ &\quad + \int_{[\Omega \setminus B(0, \varepsilon)] \cap \{z_\lambda - z_\mu + \mu - \lambda < 0\}} \frac{N-1}{|x|^2} (z_\lambda - z_\mu + \mu - \lambda)(x) dx. \end{aligned}$$

Now one has to distinguish two cases:

If  $N \geq 3$ , then, given that  $C := \|z_\lambda - z_\mu + \mu - \lambda\|_\infty < +\infty$ , one has

$$0 \leq -\frac{N-1}{\varepsilon} \int_{\partial B(0, \varepsilon) \cap \{z_\lambda - z_\mu + \mu - \lambda < 0\}} (z_\lambda - z_\mu + \mu - \lambda)(x) dx \leq 2\pi C \varepsilon^{N-2}.$$



Thus sending  $\varepsilon \rightarrow 0$  and considering identity (4.3) we obtain

$$\int_{\{z_\lambda - z_\mu + \mu - \lambda < 0\}} \frac{N-1}{|x|^2} (z_\lambda - z_\mu + \mu - \lambda)(x) dx \geq 0$$

hence

$$z_\lambda - \lambda \geq z_\mu - \mu.$$

If  $N = 2$  and  $\varepsilon$  is small then, by Lemma 4.2,

$$\partial B(0, \varepsilon) \cap \{z_\lambda - z_\mu + \mu - \lambda < 0\} = \emptyset$$

thus

$$-\frac{N-1}{\varepsilon} \int_{\partial B(0, \varepsilon) \cap \{z_\lambda - z_\mu + \mu - \lambda < 0\}} (z_\lambda - z_\mu + \mu - \lambda)(x) dx = 0$$

and the conclusion follows in the same way as above.  $\square$

**Corollary 4.4.** *For any  $\lambda, \mu \geq 0$ , one has*

$$\|z_\lambda - z_\mu\|_\infty \leq |\lambda - \mu|.$$

Another very important consequence of Proposition 4.3 is that solutions  $(u_\lambda)_{\lambda \in \mathbb{R}^+}$  form a semigroup. To prove this, let us denote

$$T_\lambda(g) = \operatorname{argmin}_{u \in BV(\Omega)} \lambda \int_\Omega |Du| + \frac{1}{2} \|u - g\|_2^2$$

for any real  $\lambda \geq 0$  and any datum  $g \in L^2(\Omega)$ . Thus,

$$T_\lambda = (I + \lambda \partial TV)^{-1}$$

is the *resolvent operator* discussed in [9].

**Proposition 4.5.** *If  $\Omega = B(0, R)$ ,  $g \in L^2(\Omega)$  is radial,  $\mu > \lambda \geq 0$  are two regularization parameters then*

$$T_\mu(g) = T_{\mu-\lambda} \circ T_\lambda(g).$$

*Proof. Step 1:* Let us set

$$z_0 = \operatorname{argmin}_{z \in A_{\mu-\lambda}} \int_\Omega (\operatorname{div}(z) + T_\lambda(g))^2$$

and  $z'_\mu = z_0 + z_\lambda$ . We claim that  $\operatorname{div}(z'_\mu) = \operatorname{div}(z_\mu)$ .

On the one hand, we know from the previous corollary that  $z_\mu - z_\lambda \in A_{\mu-\lambda}$  thus by comparison with  $z_0$

$$\int_\Omega (\operatorname{div}(z'_\mu - z_\lambda) + T_\lambda(g))^2 \leq \int_\Omega (\operatorname{div}(z_\mu - z_\lambda) + T_\lambda(g))^2.$$

and using the Euler-Lagrange equation it follows that

$$\int_\Omega (\operatorname{div}(z'_\mu) + g)^2 \leq \int_\Omega (\operatorname{div}(z_\mu) + g)^2.$$

Whereas, on the other hand,  $z'_\mu \in A_\mu$  thus by comparison with  $z_\mu$

$$\int_{\Omega} (\operatorname{div}(z_\mu) + g)^2 \leq \int_{\Omega} (\operatorname{div}(z'_\mu) + g)^2$$

which proves our claim.

*Step 2:* Now considering the following Euler-Lagrange equations

$$\begin{cases} \operatorname{div}(z_\mu) = T_\mu(g) - g, \\ \operatorname{div}(z_\lambda) = T_\lambda(g) - g, \\ \operatorname{div}(z_0) = T_{\mu-\lambda} \circ T_\lambda(g) - T_\lambda(g), \end{cases}$$

the conclusion follows readily from the result of Step 1.  $\square$

Let us recall that given an open set  $\Omega \subset \mathbb{R}^N$  and an initial condition  $u(0) = g \in L^2(\Omega)$  there exists a unique solution to the gradient flow equation (see [5]):

$$-\partial_t u(t) \in \partial TV(u(t)) \text{ a.e. in } t \in [0, T]. \quad (4.4)$$

We are also going to need the following classical result for the resolvent operator of a maximal monotone operator (see [9, Corollary 4.4]):

**Proposition 4.6.** *Let  $\Omega$  be an open set of  $\mathbb{R}^N$  and  $u(t)$  be the solution of (4.4) with an initial condition  $u(0) = g \in L^2(\Omega)$  then*

$$\lim_{n \rightarrow +\infty} T_{\frac{t}{n}}^n(u(0)) = \lim_{n \rightarrow +\infty} \left( I + \frac{t}{n} \partial TV \right)^{-n} (u(0)) = u(t),$$

the convergence taking place in  $L^2(\Omega)$ .

With these results in hands we are ready to prove the main result of this section, namely

**Theorem 4.7.** *Let  $\Omega = B(0, R)$ ,  $g \in L^2(\Omega)$  be radial and  $u(t)$  be the solution of (4.4) with an initial condition  $u(0) = g$ . Then  $u(t)$  is the unique minimizer of*

$$\min_{u \in BV(\Omega)} t \int_{\Omega} |Du| + \frac{1}{2} \|u - g\|_2^2. \quad (4.5)$$

*Proof.* From Proposition 4.5 it follows

$$T_{\frac{t}{n}}^n(u(0)) = T_t(u(0))$$

and making  $n \rightarrow +\infty$  one has by Proposition 4.6

$$u(t) = T_t(u(0))$$

hence the statement of the theorem.  $\square$

**Remark 4.8.** *This property is not true for general semi-groups  $u(t)$ . Let us reason by contradiction. If  $u(t)$  solves both (4.4) and (4.5) then writing the Euler-Lagrange equation one has for some fixed  $t > 0$ :*

$$\begin{cases} -\partial_t u(t) \in \partial TV(u(t)), \\ \frac{g - u(t)}{t} \in \partial TV(u(t)). \end{cases}$$

Placing ourselves on the set  $\{\nabla u(t) \neq 0\}$ , it follows

$$-t\partial_t u = g - u,$$

hence

$$\frac{d}{dt} \left( \frac{u}{t} \right) = -\frac{1}{t^2}(u - t\partial_t u) = -\frac{g}{t^2} = \frac{d}{dt} \left( \frac{g}{t} \right).$$

Thus, there exists  $C(x)$  that does not depend on  $t$  such that

$$u(t, x) = g(x) + C(x)t$$

whenever  $\nabla u(t) \neq 0$  in a neighborhood of  $x$ . This contradicts the example of Allard where  $\Omega = \mathbb{R}^2$  and  $g = \chi_{[0,1] \times [0,-1] \cup [-1,0] \times [0,1]}$ . Indeed let us consider the origin  $x_0 = 0$ . Then from the analysis of Allard (see section 3.2) one knows that for  $0 < t < t^*$  there is a neighborhood of  $x_0$  that is contained in  $\{\nabla u(t) \neq 0\}$ . Thus we get a contradiction if the minimizer  $u_t$  of (ROF) with parameter  $\lambda = t$  is not affine in  $t$  at  $x_0$ . This is indeed observed by a simple numerical experimentation:

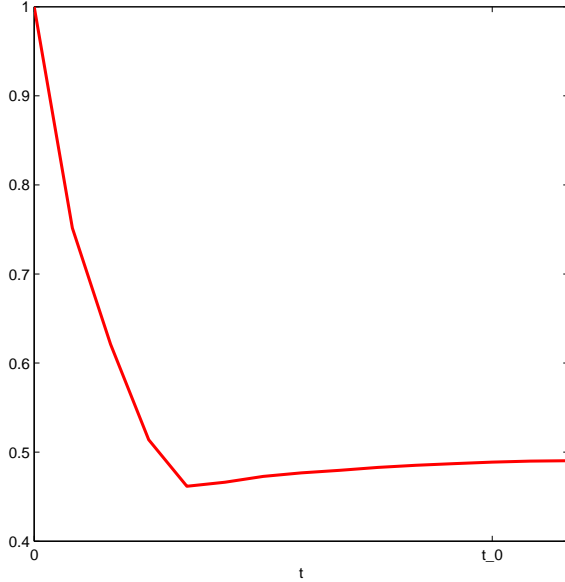


Figure 10: Evolution of  $u(t, 0)$

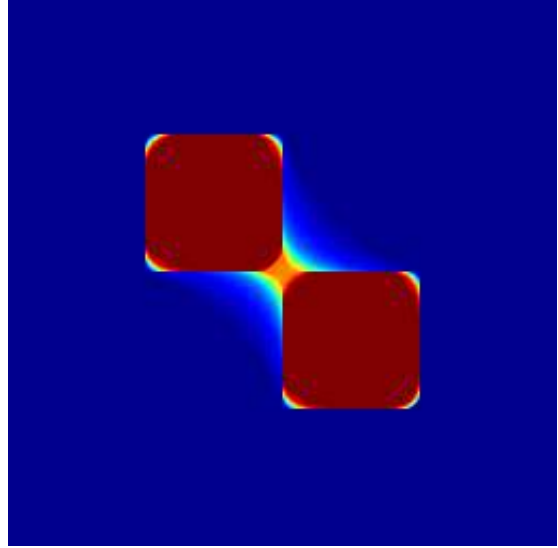


Figure 11: Solution  $u(t)$  for  $t = t_0$

The approach is rigorous since the algorithm used for the minimization of discrete (ROF) functional is convergent (see [18] and [24, Chapter 5]) and in [30] it is shown that the difference between the continuous (ROF) model and its finite difference discretization is bounded and tends to zero.

## 4.2 Staircasing and discontinuities

From Proposition 4.3 we can also get some information on the staircase regions. Indeed, whenever condition  $|z_\lambda| \leq \lambda$  is saturated, we know from the Euler-Lagrange equation that  $\nabla u_\lambda \neq 0$ . This way, we can get an inclusion principle for the staircasing namely:

**Theorem 4.9.** *For any reals  $\mu > \lambda \geq 0$ , one has  $\{|z_\lambda| < \lambda\} \subset \{|z_\mu| < \mu\}$ .*

This is not true in general as seen through the example of Allard (see Section 3.2).

**Remark 4.10.** *Theorem 4.7 is a key step in [10] to prove that in the one-dimensional case, staircasing occurs almost everywhere for a noisy 1D signal. By noisy we mean that the original signal was perturbed by a Wiener process. We cannot expect to get from our analysis such a result simply because a radial function that underwent an addition of noise is not radial anymore. We could however extend their result to the case of a radial “noise” but this does not seem really interesting. However, it would tell us again that staircasing is an important phenomenon for minimizers of perturbed signals.*

As for the discontinuities one can actually refine the results of [23] thanks to the following

**Theorem 4.11.** *Let  $\Omega = B(0, R)$  and  $g \in L^N(\Omega)$  be radial. Consider also  $\mu > \lambda > 0$  and  $u_\lambda, u_\mu$  the corresponding minimizers of (ROF). Then one has*

$$J_{u_\mu} \subset J_{u_\lambda}.$$

*If in addition,  $g \in L^N(\Omega) \cap BV(\Omega)$  then given  $\mu > \lambda \geq 0$*

$$J_{u_\mu} \subset J_{u_\lambda} \subset J_g.$$

*Proof.* Once we know Theorem 4.7 the result is a straightforward consequence of [12, Theorem 4.1] and [14, Theorem 2] which establishes a similar inclusion principle for the TV flow.  $\square$

## 5 Conclusion and perspective

In this paper, we examined some fine properties of the total variation minimization problem. In particular, we established that the staircasing phenomenon always occurs for the continuous ROF problem. We refined this result in the radial case by proving that the staircase zones are non-decreasing with the regularization parameter. The argument is based on the relation between the ROF problem and the total variation flow, which we prove to hold for radial data. In particular, using known results for the flow we also get that the discontinuities form a monotone sequence given that the datum is radial.

Most of the results our rely heavily on the connection with the perimeter problem via the coarea formula and it would interesting to understand how they can be adapted to take into account linear perturbations of the data (convolution but also Radon or Fourier transforms).

An interesting extension of our work would be to prove an  $N$ -dimensional counterpart for the result of the recent paper [10]: staircasing occurs almost everywhere for noisy images.

## References

- [1] W. K. ALLARD, *Total variation regularization for image denoising. II. Examples*, SIAM J. Imaging Sci., 1 (2008), pp. 400–417.

- [2] ———, *Total variation regularization for image denoising. III. Examples*, SIAM J. Imaging Sci., 2 (2009), pp. 532–568.
- [3] F. ALTER, V. CASELLES, AND A. CHAMBOLLE, *A characterization of convex calibrable sets in  $\mathbb{R}^N$* , Math. Ann., 332 (2005), pp. 329–366.
- [4] L. AMBROSIO, N. FUSCO, AND D. PALLARA, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 2000.
- [5] F. ANDREU-VAILLO, V. CASELLES, AND J. M. MAZÓN, *Parabolic quasilinear equations minimizing linear growth functionals*, vol. 223 of Progress in Mathematics, Birkhäuser Verlag, Basel, 2004.
- [6] G. ANZELLOTTI, *Pairings between measures and bounded functions and compensated compactness*, Ann. Mat. Pura Appl. (4), 135 (1983), pp. 293–318 (1984).
- [7] M. BONFORTE AND A. FIGALLI, *Total variation flow and sign fast diffusion in one dimension*, Journal of Differential Equations, 252 (2012), pp. 4455–4480.
- [8] J. BOURGAIN AND H. BREZIS, *On the equation  $\operatorname{div} Y = f$  and application to control of phases*, J. Amer. Math. Soc., 16 (2003), pp. 393–426 (electronic).
- [9] H. BREZIS, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Publishing Co., Amsterdam, 1973. North-Holland Mathematics Studies, No. 5. Notas de Matemática (50).
- [10] A. BRIANI, A. CHAMBOLLE, M. NOVAGA, AND G. ORLANDI, *On the gradient flow of a one-homogeneous functional*, Confluentes Mathematici, 3 (2011), pp. 617–635.
- [11] L. A. CAFFARELLI AND R. DE LA LLAVE, *Planelike minimizers in periodic media*, Comm. Pure Appl. Math., 54 (2001), pp. 1403–1441.
- [12] V. CASELLES, A. CHAMBOLLE, AND M. NOVAGA, *The discontinuity set of solutions of the TV denoising problem and some extensions*, Multiscale Model. Simul., 6 (2007), pp. 879–894.
- [13] ———, *Regularity for solutions of the total variation denoising problem*, Revista Matemática Iberoamericana, 27 (2010), pp. 233–252.
- [14] V. CASELLES, K. JALALZAI, AND M. NOVAGA, *On the jump set of solutions of the total variation flow*, (2012).
- [15] A. CHAMBOLLE, *An algorithm for total variation minimization and applications*, J. Math. Imaging Vision, 20 (2004), pp. 89–97. Special issue on mathematics and image analysis.
- [16] A. CHAMBOLLE, V. CASELLES, D. CREMERS, M. NOVAGA, AND T. POCK, *An introduction to total variation for image analysis*, in Theoretical foundations and numerical methods for sparse recovery, vol. 9 of Radon Ser. Comput. Appl. Math., Walter de Gruyter, Berlin, 2010, pp. 263–340.
- [17] A. CHAMBOLLE AND P.-L. LIONS, *Image recovery via total variation minimization and related problems*, Numer. Math., 76 (1997), pp. 167–188.

- [18] A. CHAMBOLLE AND T. POCK, *A first-order primal-dual algorithm for convex problems with applications to imaging*, J. Math. Imaging Vision, 40 (2011), pp. 120–145.
- [19] A. CHAMBOLLE AND G. THOUROUDE, *Homogenization of interfacial energies and construction of plane-like minimizers in periodic media through a cell problem*, Netw. Heterog. Media, 4 (2009), pp. 127–152.
- [20] T. DE PAUW AND W. PFEFFER, *Distributions for which  $\operatorname{div} v = f$  has a continuous solution*, Communications on Pure and Applied Mathematics, 61 (2008), pp. 230–260.
- [21] E. GIUSTI, *Minimal surfaces and functions of bounded variation*, vol. 80 of Monographs in Mathematics, Birkhäuser Verlag, Basel, 1984.
- [22] L. HÖRMANDER, *The analysis of linear partial differential operators. I*, Classics in Mathematics, Springer-Verlag, Berlin, 2003. Distribution theory and Fourier analysis, Reprint of the second (1990) edition [Springer, Berlin; MR1065993 (91m:35001a)].
- [23] K. JALALZAI, *Discontinuities of the minimizers of the weighted or anisotropic total variation for image reconstruction*, (2012).
- [24] ———, *Regularization of inverse problems in image processing*, PhD Thesis, École Polytechnique, Palaiseau, (March 2012).
- [25] K. JALALZAI AND A. CHAMBOLLE, *Enhancement of blurred and noisy images based on an original variant of the total variation*, Scale Space and Variational Methods in Computer Vision, (June 2009), pp. 368–376.
- [26] C. LOUCHET AND L. MOISAN, *Total variation denoising using posterior expectation*, (2008).
- [27] M. NIKOLOVA, *Local strong homogeneity of a regularized estimator*, SIAM J. Appl. Math., 61 (2000), pp. 633–658 (electronic).
- [28] W. RING, *Structural properties of solutions to total variation regularization problems*, M2AN Math. Model. Numer. Anal., 34 (2000), pp. 799–810.
- [29] L. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Physica D: Nonlinear Phenomena, 60 (1992), pp. 259–268.
- [30] J. WANG AND B. J. LUCIER, *Error bounds for finite-difference methods for Rudin-Osher-Fatemi image smoothing*, SIAM J. Numer. Anal., 49 (2011), pp. 845–868.
- [31] W. P. ZIEMER, *Weakly differentiable functions*, vol. 120 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1989. Sobolev spaces and functions of bounded variation.

